Overview of VACET Software Engineering Activities

Program Review

Rockville, MD

4/28/09

Hank Childs

VACET Chief Software Engineer

UCRL-PRES-412682

Work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

Overview of VACET Software Engineering Group (SEG)

- Goal: Deliver production quality, parallel capable software infrastructure that enables science at the petascale.
- Strategy: adapt, extend, and create software when necessary
- The SEG comes from all five VACET sites and is the focal point for many collaborations.

VACET

Our software delivery strategy is designed to maximize impact.

- We favor tool and library development, because these approaches lead to the greatest impact.
  - Developing many “one-offs” is both beyond our budget and also not a cost effective use of money.
- VisIt is our supported software for production quality petascale-capable visualization and analysis.
- Of course, a heavyweight tool is not always the best way to impact stakeholders and we adapt approaches as necessary.

U.S. DEPARTMENT OF ENERGY | Office of Science

SciDAC
Scientific Discovery through Advanced Computing

VACET

Outline

- Non-VisIt highlights
- Overview of VisIt
- VisIt highlights

U.S. DEPARTMENT OF ENERGY | Office of Science

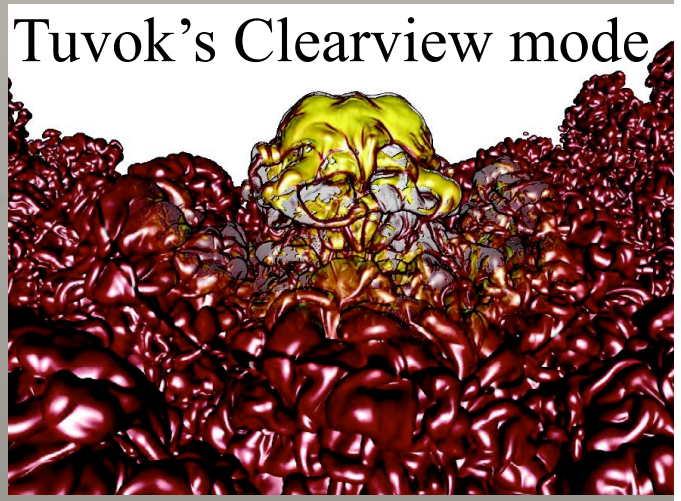
SciDAC
Scientific Discovery through Advanced Computing

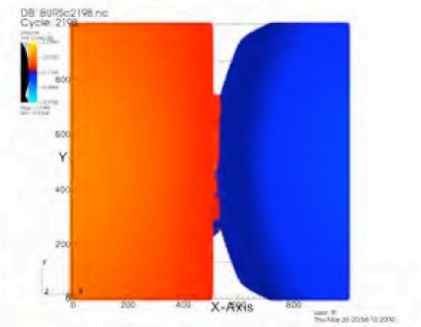
Outline

- Non-VisIt highlights
- Overview of VisIt
- VisIt highlights

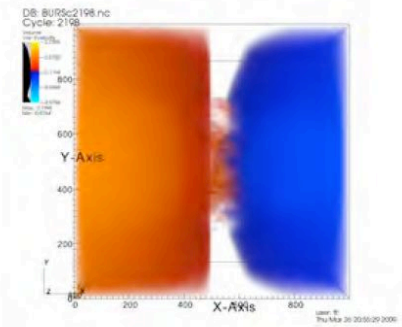
Tech transfer: the SEG is productizing the University of Utah's volume renderer.

- Utah's pioneering work in multi-dimensional transfer functions is being productized to reach a wider audience.
  - SLIVR library and its next generation replacement, Tuvok.
  - SLIVR is integrated into VisIt; Tuvok integration planned.

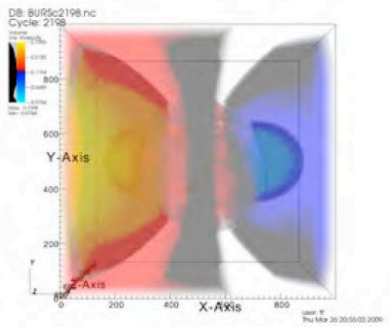




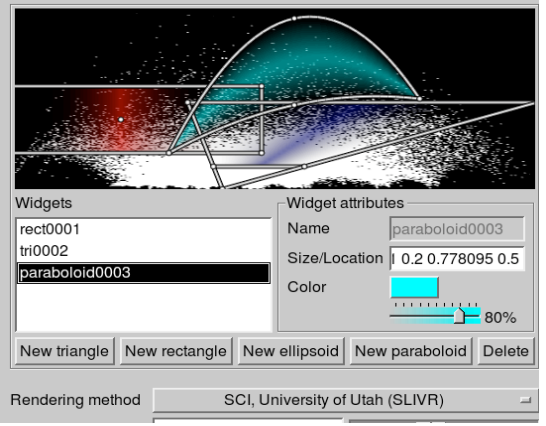
(a) Splat-based volume rendering.



(b) 3D texture-based volume rendering.

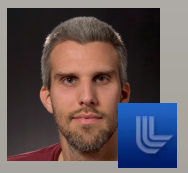


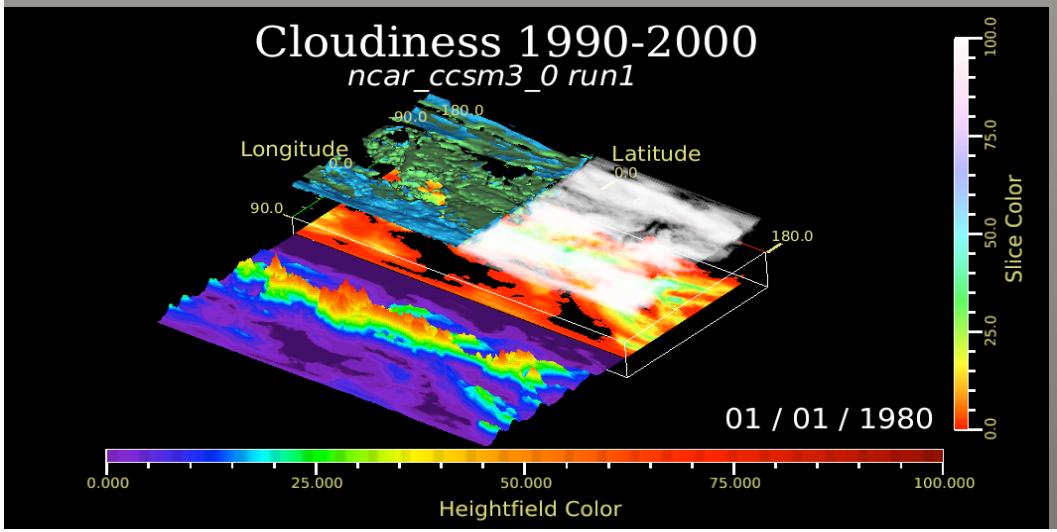
(c) SLIVR volume rendering.

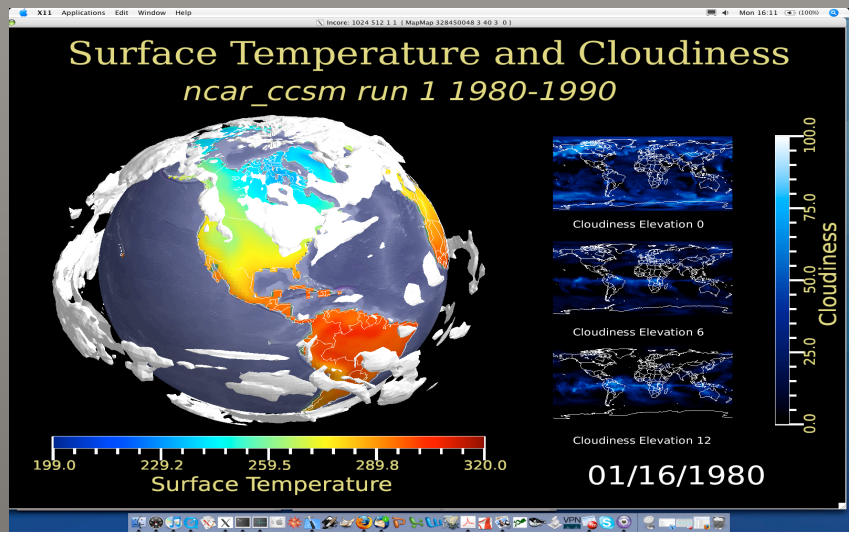


The SEG developed and integrated lightweight visualization software into CDAT.

- CDAT: dominant, highly domain specific visualization application for climate data
  - But CDAT lacked 3D capabilities.
  - Need: very lightweight library for tight integration.
- The SEG developed SFSG (Simple and Flexible Scene Graph) library and integrated and deployed it in CDAT.





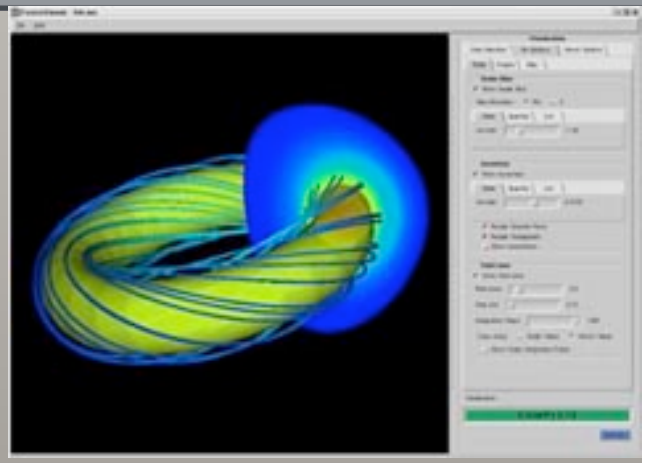


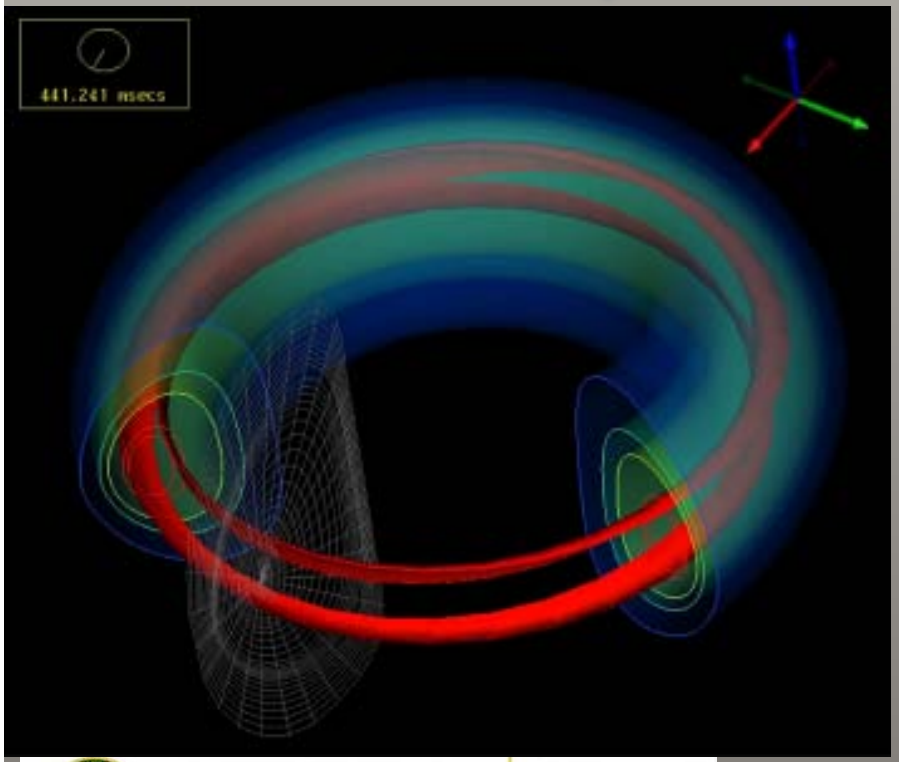
VACET

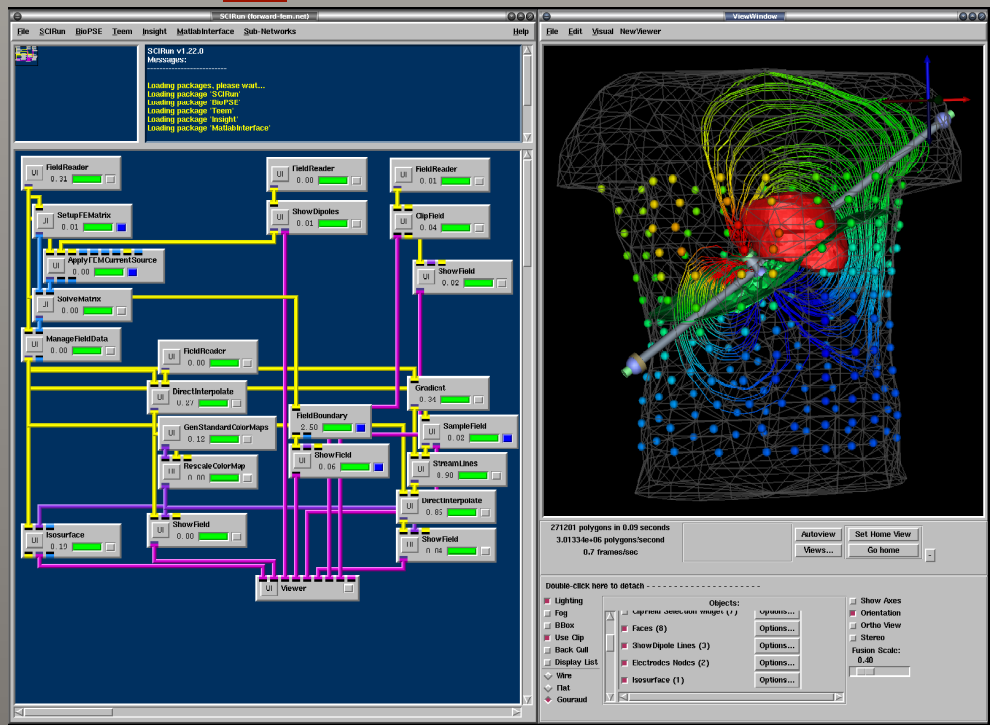
VACET Software Tools

SCIRun - www.sci.utah.edu

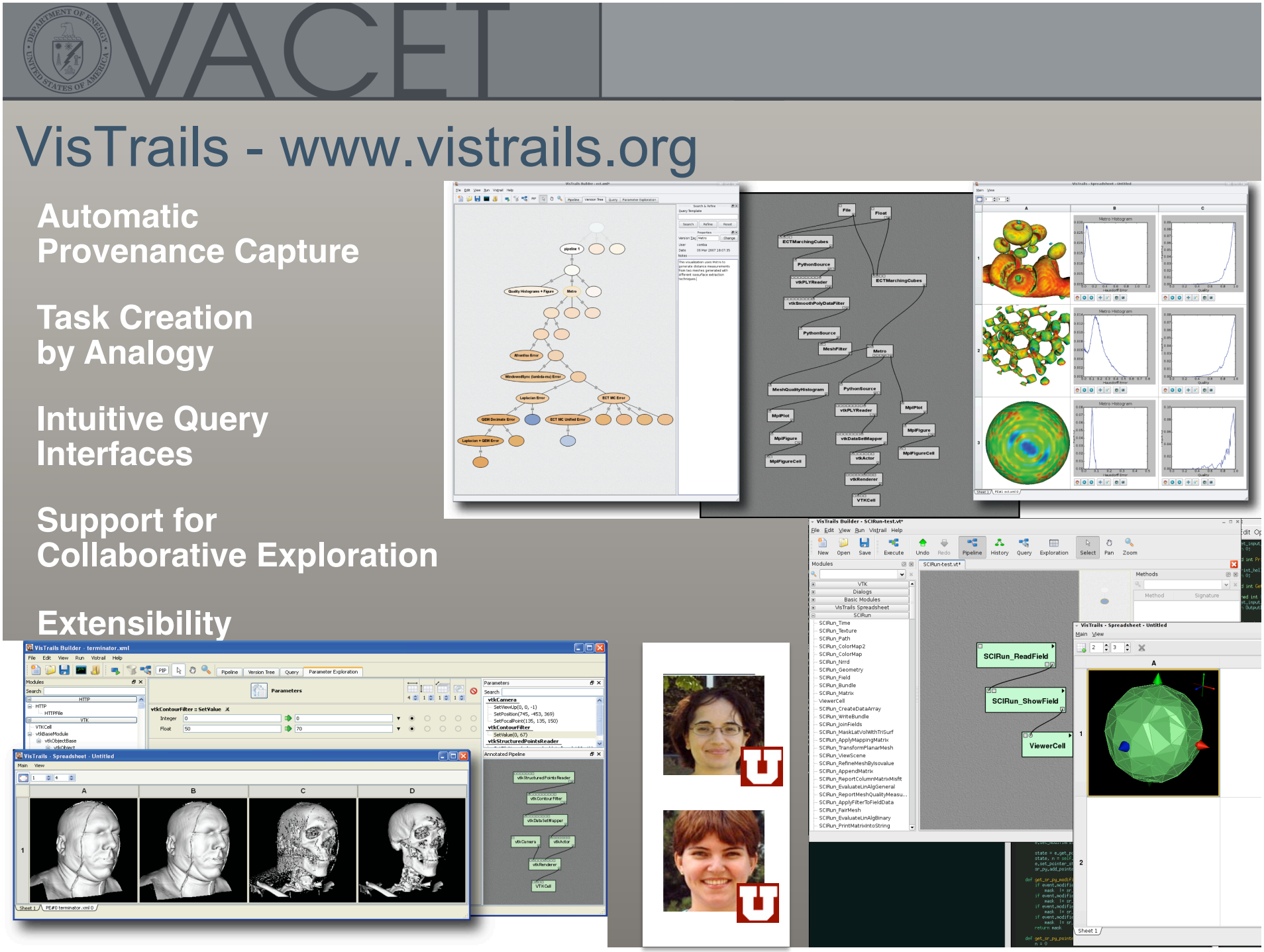
– SCI Institute, University of Utah







U.S. DEPARTMENT OF ENERGY | Office of Science


VACET
VisTrails - www.vistrails.org
Automatic
Provenance Capture
Task Creation
by Analogy
Intuitive Query
Interfaces
Support for
Collaborative Exploration
Extensibility
SCIRun_ReadField
SCIRun_ShowField
ViewerCell

VACET

Outline

- Non-VisIt highlights
- Overview of VisIt
- VisIt highlights

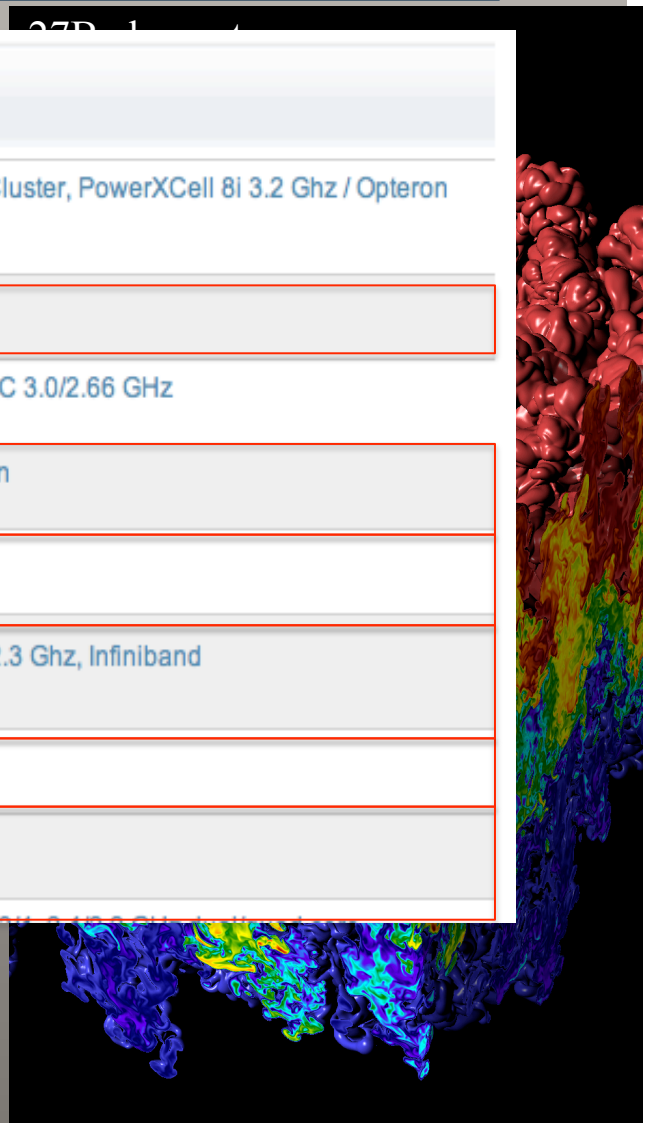
U.S. DEPARTMENT OF ENERGY | Office of Science

SciDAC
Scientific Discovery through Advanced Computing

VisIt is a richly featured, turnkey application

- Vis
  vis
  sin
  - –
  - –
- R&
- Us
  of

Rank	Site	Computer
1	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband IBM
2	Oak Ridge National Laboratory United States	Jaguar - Cray XT5 QC 2.3 GHz Cray Inc.
3	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0/2.66 GHz SGI
4	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution IBM
5	Argonne National Laboratory United States	Blue Gene/P Solution IBM
6	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband Sun Microsystems
7	NERSC/LBNL United States	Franklin - Cray XT4 QuadCore 2.3 GHz Cray Inc.
8	Oak Ridge National Laboratory United States	Jaguar - Cray XT4 QuadCore 2.1 GHz Cray Inc.

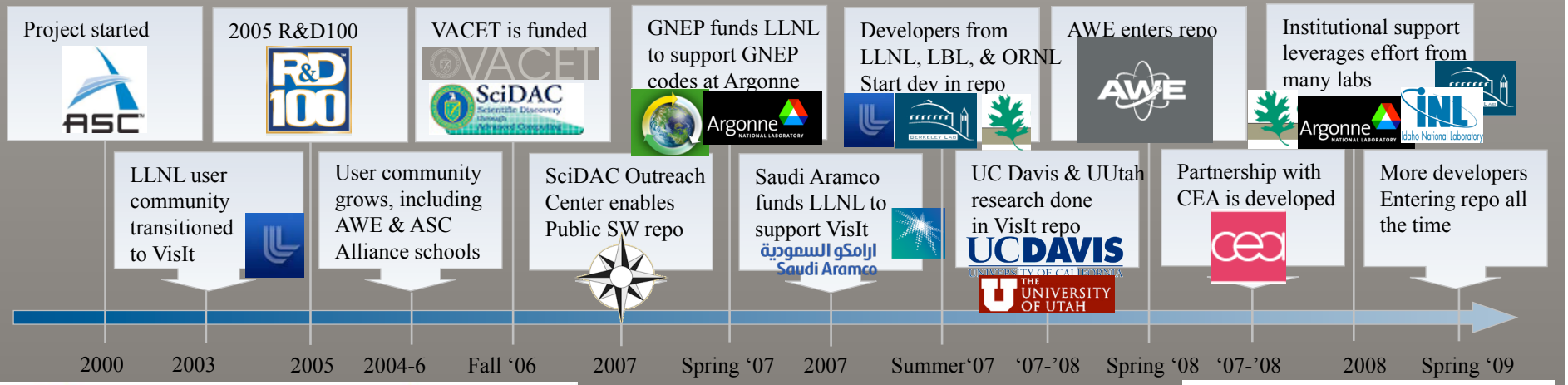


VisIt has a rich feature set that can impact many science areas.

- Meshes: rectilinear, curvilinear, unstructured, point, AMR
- Data: scalar, vector, tensor, material, species
- Dimension: 1D, 2D, 3D, time varying
- Rendering (~15): pseudocolor, volume rendering, hedgehogs, glyphs, mesh lines, etc…
- Data manipulation (~40): slicing, contouring, clipping, thresholding, restrict to box, reflect, project, revolve, …
- File formats (~85)
- Derived quantities: >100 interoperable building blocks
  +,-,*,/, gradient, mesh quality, if-then-else, and, or, not
- Many general features: position lights, make movie, etc
- Queries (~50): ways to pull out quantitative information, debugging, comparative analysis

The VisIt project provides a substantial leverage for VACET's work.

- VisIt represented over 50 person-years of effort at the time of VACET being funded
  - VisIt contains over one million lines of code
- Partnership between: Department of Energy's Office of Science, National Nuclear Security Agency, and Office of Nuclear Energy, among others



U.S. DEPARTMENT OF ENERGY | Office of Science

SciDAC Scientific Discovery through Advanced Computing

Reviewer questions about VisIt

- - How do you envision VisIt will be supported and managed after the term of the current center expires? Does it have financial support apart from SciDAC?
- - Will the scientists who depend on VisIt continue to require a team of dedicated software engineers to make updates, or will the software eventually be transitioned to a community-supported, open source type of product?
- - Do you have any indication of how much traction VisIt has gained in the broader scientific community, outside of the group of people with whom you directly collaborate? For example, do you know how many times it has been downloaded from your website or how often it has been credited in papers published by others?

Outline

- Non-VisIt highlights
- Overview of VisIt
- VisIt highlights

The role of VACET in VisIt.

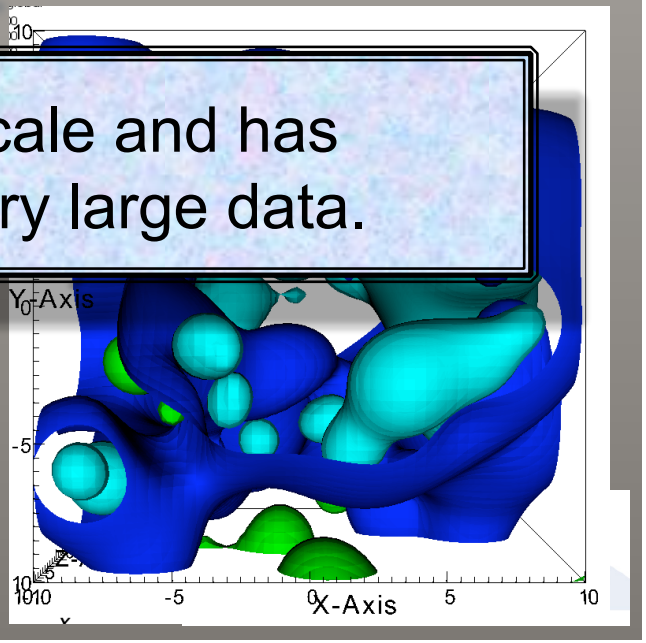
- Everything in the rest of this presentation describes work done by VACET.
  - (A fraction of the work done by VACET!)
- Eric Brugger, NNSA VisIt lead: "VACET has assumed the lead role in ensuring that VisIt will be able to handle the largest data sets."
- NNSA has reduced resources on VisIt from 6 FTEs in 2006 to 4 FTEs today.
  - The NNSA VisIt team's is focusing on user support for NNSA customers.

VisIt as a petascale visualization and data analysis engine.

- VisIt is undergoing Joule code certification
  - VisIt will be the first ever non-simulation Joule code
- As part of the certification process, we demonstrated VisIt on a one trillion cell data set.
  - 20K cores and 40K cores on jaguar.

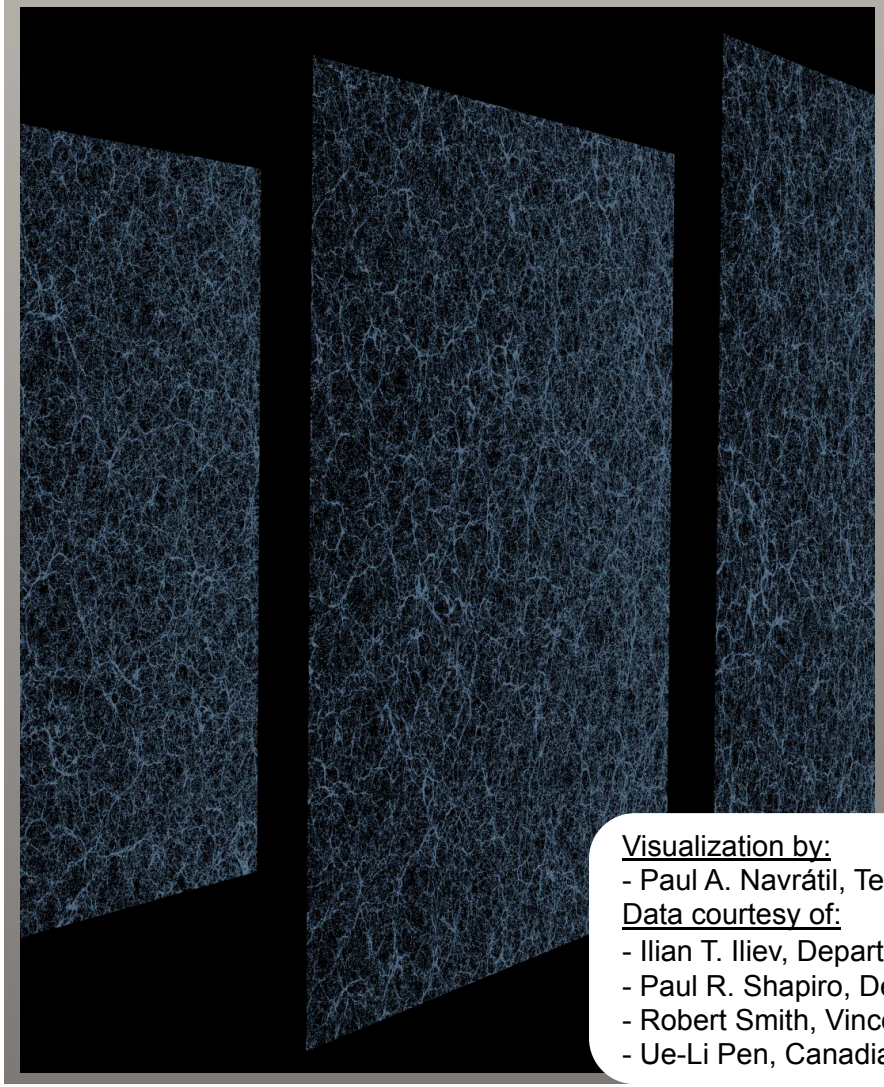
The SEG ensures our tools will scale and has demonstrated good success on very large data.

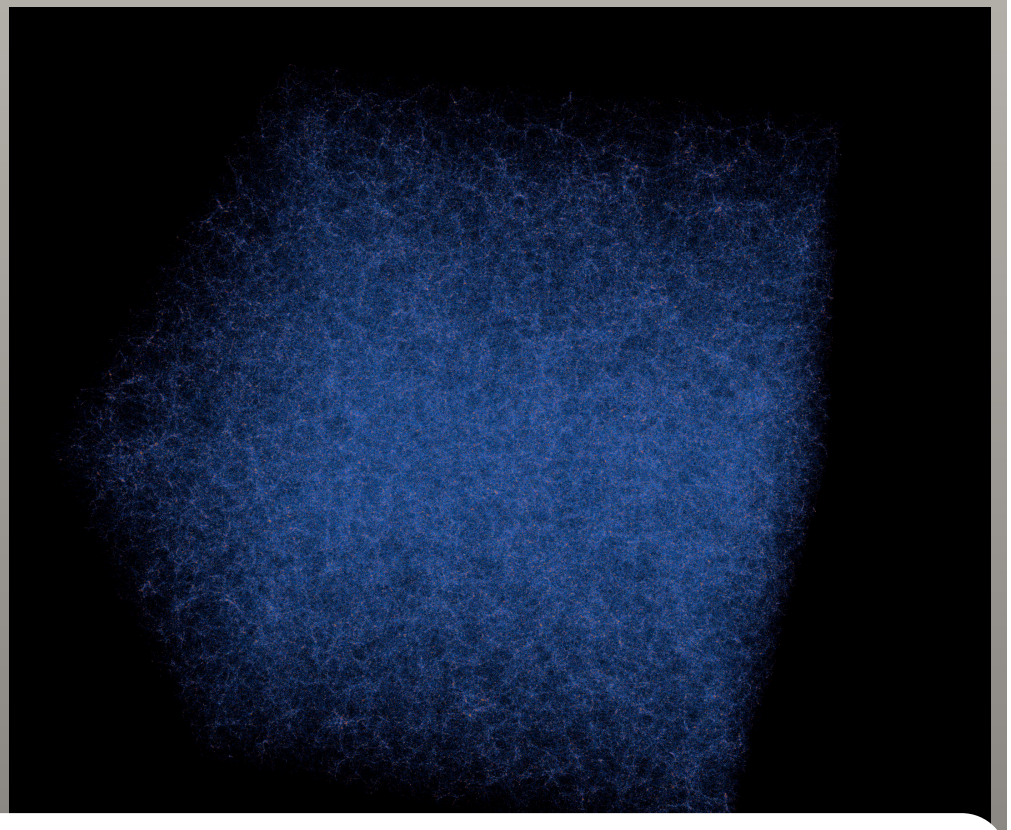
machines.



Joule code certification

VisIt also has been demonstrated on a 216B cell “real” data set.





Visualization by:
- Paul A. Navrátil, Texas Advanced Computing Center, the University of Texas at Austin

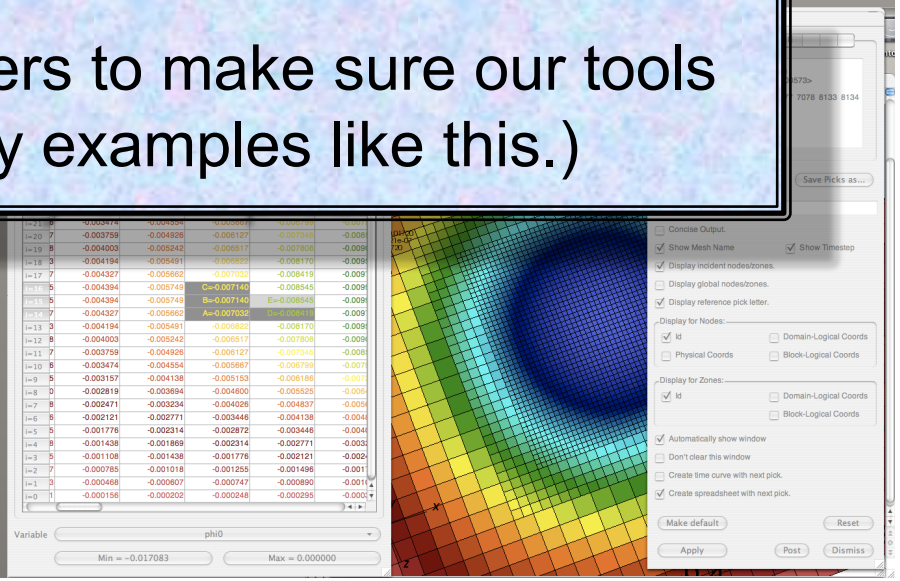
Data courtesy of:
- Ilian T. Iliev, Department of Physics and Astronomy, University of Sussex, U.K.
- Paul R. Shapiro, Department of Astronomy, the University of Texas at Austin
- Robert Smith, Vincent Desjacques, Institute for Theoretical Physics, University of Zürich, Switz.
- Ue-Li Pen, Canadian Institute for Theoretical Astrophysics, University of Toronto, Canada

VACET has improved VisIt's AMR capabilities significantly.

- Improvements:
  - Optimized many, many routines, resulting in end-to-end speedups of >10X
  - Added critical debugging functionality, such as
  - Macros and python callbacks for customizability.
  - Responded to over 30 requests for interface changes and bug fixes.

The SEG works with customers to make sure our tools will fit their needs. (Many examples like this.)

- Done for collaborators at APDEC & CAC.



VACET

We are using FastBit to accelerate exploration of Laser Wakefield simulations

- Allow for interactive seeking of particles, with execution time linear in the amount of returned particles.
- Collaboration with the SDM center.
- SC08 paper on this topic.

The SEG transitions research done by VACET into VACET's tools.






SciDAC
Scientific Discovery through Advanced Computing

VACET

Streamlines represent an activity spanning research to SWE to data insight.

Customers : "VisIt is bad at streamlines"

"We need parallel streamlines" / "Wow! Parallel streamlines is hard"

Research effort: Efficient parallel streamline generation



UC DAVIS
UNIVERSITY OF CALIFORNIA

Scalable Computation of Streamlines and Pathlines on Very Large Datasets *

Dave Pugmire, Oak Ridge National Laboratory, PO Box 2008 MS6016, Oak Ridge, TN 37831-6016, pugmire@ornl.gov

Christoph Garth, University of California, Davis, One Shields Road, Davis, CA, 95616, cgarth@ucdavis.edu

Hank Childs, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA, 94550, childs3@llnl.gov

Sean Ahern, Oak Ridge National Laboratory, PO Box 2008 MS6016, Oak Ridge, TN 37831-6016, ahern@ornl.gov

Gunther Weber, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA, 94701, ghweber@lbl.gov

ABSTRACT

Understanding vector fields resulting from large scientific simulations is an important and often difficult task. Streamlines, i.e. curves that are tangential to a given vector field at each point, and pathlines, the time varying analog of streamlines, are a convenient and powerful visualization method in this context. More recently, techniques for analysis and visualization of vector fields that make use of streamlines as a building block for further knowledge discovery have added to the importance of reliable and efficient streamline computation. For very large data and massive numbers of streamlines, efficient streamline computation poses unique challenges in balancing the computational demands placed on I/O, memory, communication, and processors within an HPC system. This is a consequence of the highly non-linear, non-local, and data-dependent nature of streamline computation. In this paper we present several scalable algorithms for computing streamlines, discuss the respective distribution strategies, and illustrate and compare performance and efficiency results for each algorithm.

Keywords

visualization, streamlines, pathlines, scaling, parallel

1. INTRODUCTION

*(Does NOT produce the permission block, copyright information nor page numbering). For use with ACM_PROC_ARTICLE-SP.CLS. Supported by ACM.

Christoph TODO

1. Motivate SLs
2. Motivate Parallel SLs
3. Paper outline

The paper is a shoot-out between different streamline and pathline parallelization/distribution variants. Overall goal is to demonstrate three different parallelization strategies (Load On Demand, Static, Hybrid), document their absolute and relative performance in both real-world and synthetic cases, and identify guidelines that help choosing one of these strategies for a given problem.

Parameters for such guidelines are (restricted to a binary classification for the first three in the interest of feasibility):

1. dataset size (small/large)
2. seed set size (small/large)
3. seed set distribution (uniform in domain/concentrated around point of interest)
4. vector field complexity (simple / complex, where the meaning of these terms needs to be defined)

This gives 16 possible scenarios, which is hopefully enough to formulate our guidelines.

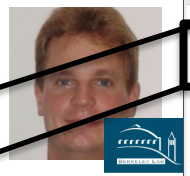
We can identify the following prototypical application scenarios that cover a wide range of typical problems requiring streamline/pathline computation:

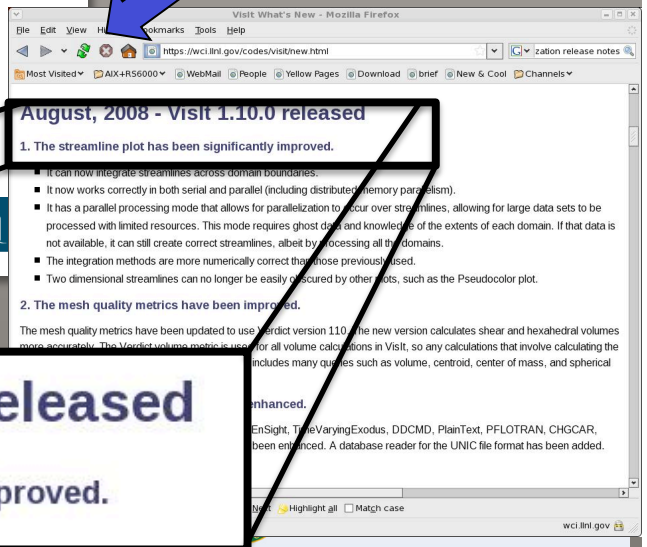
- fusion data, poincare plots, streamlines (small data, small seed set, uniformly distributed)

SC Portland, OR 2009

Deployment effort in VisIt by SEG







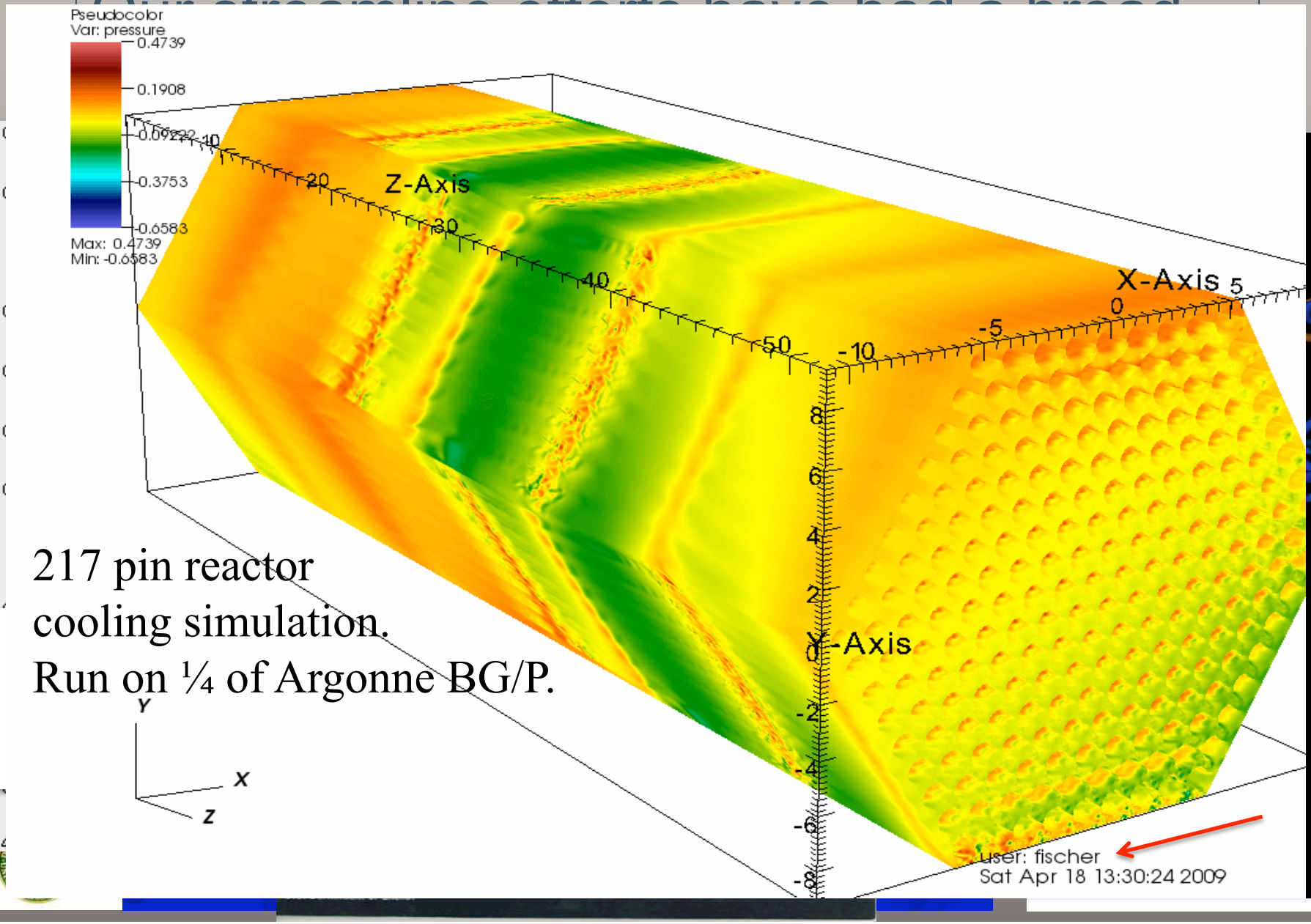
August, 2008 - VisIt 1.10.0 released

1. The streamline plot has been significantly improved.

U.S. DEPARTMENT OF ENERGY

VACET

Our streamline efforts have had a broad



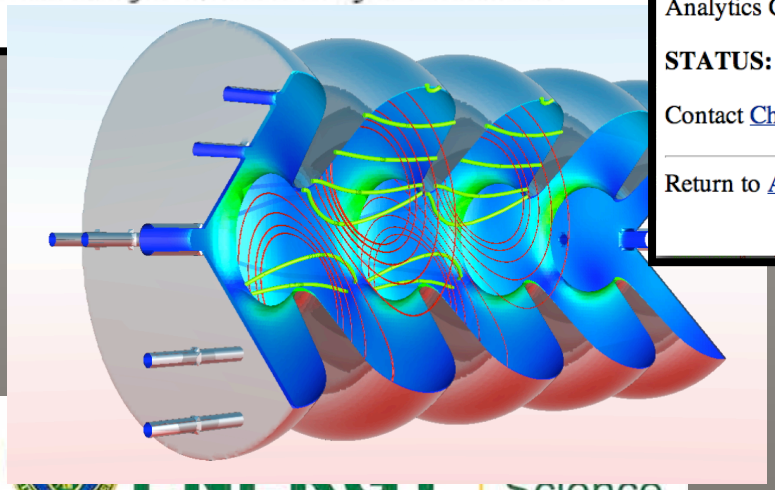
217 pin reactor
cooling simulation.
Run on ¼ of Argonne BG/P.

Our efforts have led to VisIt being adopted by important ASCR stakeholders.

In summary, I am sold on VisIt and the VACET team. They are providing a great service to our community in not only providing a tool, but also in providing usable tool that helps scientist extract knowledge from the huge amounts of data that are being generated by SciDAC applications now in use.

Sincerely,

John R. Cary
CEO, Tech-X
Prof. Physics, University of Colorado



ChomboVis

http://seesar.lbl.gov/anag/chombo/chombovis.html

apdec visit chombovis

VisIt-tutori... Center Wiki MyLLNL – Portal Apple (123) Amazon eBay Yahoo! News (1061)

ChomboVis

VisIt has superseded ChomboVis as the visualization and analysis tool of choice to Chombo (AMR) HDF5 datasets. ChomboVis is no longer being maintained or developed. Chombo users are strongly encouraged to download, install, and use VisIt. Development of VisIt for Chombo AMR datasets is ongoing and future releases of VisIt should provide addition functionality for Chombo AMR datasets.

The development of VisIt for Chombo datasets is being done by VACET (Visualization and Analytics Center for Enabling Technologies).

STATUS: No longer being maintained or developed (see above).

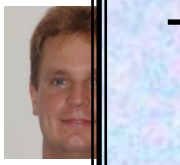
Contact ChomboVis Development Team

Return to ANAG Home Page

+ see our review document for 9 more letters of support from SciDAC-funded groups using VisIt

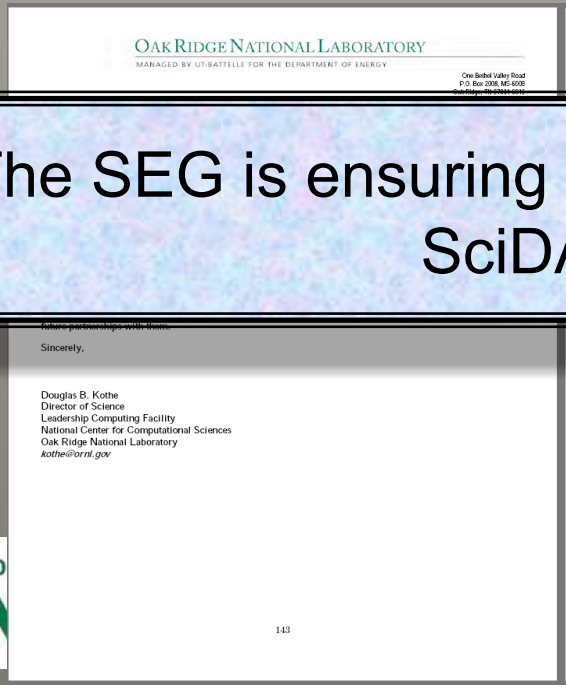
We have ensured that our tools will run on the Office of Science's big machines



- Porting for Franklin (LBL), Jaguar (ORNL), and Eureka (Argonne, connected to BG/P)
- Changes to ensure job launching, client-server connections, I/O, etc run smoothly

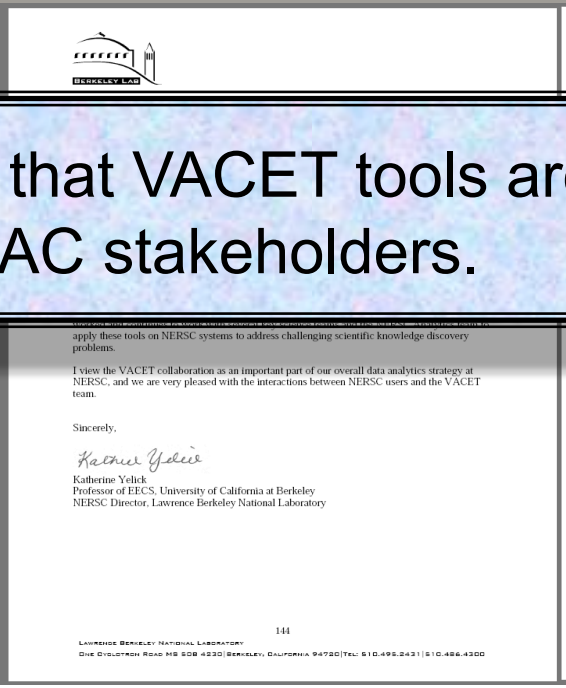
The SEG is ensuring that VACET tools are available to SciDAC stakeholders.


OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Sincerely,

Douglas B. Kothe
Director of Science
Leadership Computing Facility
National Center for Computational Sciences
Oak Ridge National Laboratory
kothe@ornl.gov

143


BERKELEY LAB

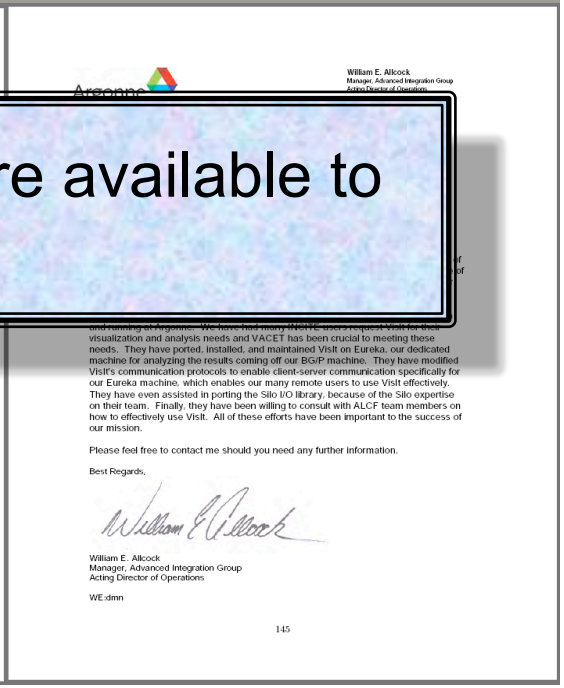
I view the VACET collaboration as an important part of our overall data analytics strategy at NERSC, and we are very pleased with the interactions between NERSC users and the VACET team.

Sincerely,

Katherine Yelick
Professor of EECS, University of California at Berkeley
NERSC Director, Lawrence Berkeley National Laboratory

144

LAWRENCE BERKELEY NATIONAL LABORATORY
ONE CYCLOTRON ROAD MS 50B 4230|BERKELEY, CALIFORNIA 94720|TEL: 510.495.2431|510.486.4300


William E. Allcock
Manager, Advanced Integration Group
Acting Director of Operations

visualization and analysis needs and VACET has been crucial to meeting these needs. They have ported, installed, and maintained Visit on Eureka, our dedicated machine for analyzing the results coming off our BG/P machine. They have modified Visit's communication protocols to enable client-server communication specifically for our Eureka machine, which enables our many remote users to use Visit effectively. They have even assisted in porting the Silo I/O library, because of the Silo expertise on their team. Finally, they have been willing to consult with ALCF team members on how to effectively use Visit. All of these efforts have been important to the success of our mission.

Please feel free to contact me should you need any further information.

Best Regards,

William E. Allcock
Manager, Advanced Integration Group
Acting Director of Operations

WE:dmn

145

We have contributed more than just algorithms to VisIt.

- Production software: bug tracking, regression testing, regular releases
- Through our efforts, VisIt now feels much more like an open source project:
  - Transitioned software repository from being "trapped at LLNL" to publicly accessible.
    - ~25 developers with write access from 9 institutions
  - Public mailing lists, archived and searchable.
    - ~300 subscribers, get ~300 posts per month
  - Wiki pages on usage and development.






Future work for the SEG: lots!.

- Improvements for existing customers
  - Many requests ranging from minor bug fixes to major new functionality.
- Additional partnerships
- Deploy research (e.g. topology, uncertainty vis, volume rendering, and integral curves)
- Implement and deploy advanced data processing algorithms for petascale data.
  - Multi-resolution data exploration, for AMR data and also for Tuvok volume rendering.
  - In situ

VACET SEG Summary

- The VACET SEG delivers production quality, parallel capable software infrastructure that enables science at the petascale.
  - We perform any and all tasks necessary to accomplish this goal.
- The SGE adapts its strategy based on need:
  - library development
  - VisIt development
- The SEG comes from all five VACET sites and is the focal point for many collaborations.
- Lots of work still to do!